

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 11-338647

(43)Date of publication of application : 10.12.1999

(51)Int.Cl.

G06F 3/06

(21)Application number : 10-142571

(71)Applicant : HITACHI LTD

(22)Date of filing : 25.05.1998

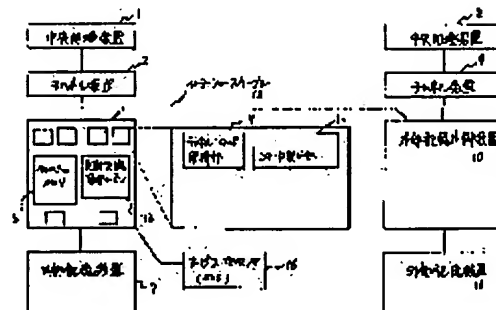
(72)Inventor : SUISHU KAZUTO
NAGASAWA MITSUO
KIMURA YUKIHISA
KOIDE TAKESHI

(54) METHOD FOR GENERATING DUPLEXED DATA

(57)Abstract:

PROBLEM TO BE SOLVED: To duplex data with high reliability while minimizing decrease in information processing capability by making a choice whether data of a main system and a subordinate system are doubled synchronously or asynchronously according to characteristics, etc., of the data.

SOLUTION: An external storage control unit is provided with an update change management table 13 which manages information whether the data of the main system and subordinate system are duplexed synchronously or asynchronously, and a service processor 16 is connected which can set a change to update data to an external storage device 11 of the subordinate system, etc. A channel processing processor is provided with a channel command analysis part 14 and a command issue part 15. When data which is important to a user are updated, the consistency between the data of both the main and subordinate systems is given importance, so the synchronous duplexing is set with high possibility, but when data whose response is taken into consideration are updated, the asynchronous duplexing is set with high possibility.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

* NOTICES *

Japan Patent Office is not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. **** shows the word which can not be translated.
3. In the drawings, any words are not translated.

CLAIMS

[Claim(s)]

[Claim 1] External storage and an external memory control unit which controls this (let these be positive side systems collectively hereafter), Another external memory control unit subordinate's external storage which is the data processing system which consists of high order equipment, and was connected in an information communication link of arbitration A formation method of doubleness data characterized by the ability to update information on a sub** system to a synchronization or asynchronous for every unit of data in data duplex system which sets (to let these be sub** systems collectively hereafter), and holds the same data as backup data.

[Claim 2] It is the formation method of doubleness data characterized by renewal of information on a sub** system enabling setting out of a synchronization or carrying out by being asynchronous for every unit of data in claim 1 at arbitration.

[Translation done.]

* NOTICES *

Japan Patent Office is not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. **** shows the word which can not be translated.
3. In the drawings, any words are not translated.

DETAILED DESCRIPTION

[Detailed Description of the Invention]

[0001]

[The technical field to which invention belongs] About the formation method and external memory control unit of doubleness data, this invention is holding the same data at details in the external storage of the inside of the external storage of a positive side system, and a sub** system, and attaining doubleness of data, and relates to data transfer technology especially about the technology which raises the reliability of data, and integrity.

[0002]

[Description of the Prior Art] It is common to hold the data used with high order equipment to external storage, to prepare a different external memory control unit subordinate the external storage which has the same data as a method of doubling data, and to connect between that each memory control unit of the exterior in the information communication link of arbitration.

[0003] Like, when [in the well-known example ***** No. 509565 / eight to / official report] a write-in instruction is published from high order equipment under such system environment, there is technology of attaining data doubleness by publishing the same write-in instruction to the external memory control unit of the secondary-data storage system by which the external memory control unit of a primary data storage system is connected with this. It is the technology of publishing the same write-in instruction as the external memory control unit of a secondary-data storage system synchronizing with the time of the write-in instruction receipt from high order equipment, and since the following write-in instruction is not received until the renewal of data of a secondary-data storage system is completed by this method, specifically, doubleness of data is certainly realizable. However, since a completion report to high order equipment is kept waiting until write-in processing of a secondary-data storage system is completed, the response to high order equipment becomes large, and cannot be said to be an efficient system.

[0004] Moreover, since such system environment has a high possibility of being comparatively installed in a remote place for the purpose of suppressing the data loss by disaster, such as an earthquake and a fire, to the minimum, it will have big effect also on a system throughput.

[0005] On the other hand, with the time of the write-in instruction receipt from high order equipment, the technology of publishing a write-in instruction to the external memory control unit of a secondary-data storage system is in asynchronous. By this method, since it can carry out by not concerning the write-in instruction issuance to the external memory control unit of a secondary-data storage system with enjoyment of data with high order equipment, response time becomes early, but since renewal of data of a secondary-data storage system is performed to asynchronous, the time zone which cannot take the consistency of the data of a primary data storage system and a secondary-data storage system exists. For this reason, the recovery of the data at the time of disaster becomes complicated.

[0006] Furthermore, the two above-mentioned means must choose either beforehand. That is, inefficient processing in which it must dare perform the data which may update by asynchronous process by synchronous processing will occur.

[0007]

[Problem(s) to be Solved by the Invention] It is asynchronous and the object of this invention is to offer the technology which chooses whether doubleness of the data of a positive side system and a sub** system is synchronously performed according to the property and significance of data to double, or it carries out, and doubles data efficiently.

[0008] Other objects of this invention are to offer the technology of realizing doubleness of reliable data, suppressing lowering of the original information processing capacity which a system has to the minimum.

[0009]

[Means for Solving the Problem] The object of this invention is attained by a means by which an external memory control unit of a positive side system recognizes (1) synchronization and asynchronous, and means to set up whether doubleness of data of a positive side system and a sub** system is synchronously performed per a storage device of arbitration, and information storage through (2) high-order equipment or a service processor, or it carries out by being asynchronous.

[0010]

[Embodiment of the Invention] Hereafter, the gestalt of one example of this invention is explained, referring to a drawing.

[0011] Drawing 1 is the block diagram of the data processing system which is the gestalt of one example of this invention. External storage 7 connected with a central processing unit 1, a channel unit 2, and the external memory control unit 3 at this is consisted of by the positive side system. Moreover, the external memory control unit 3 contains the channel processing activation processor 4 which exchanges data with a channel unit 2, cache memory 5, and the drive processing activation processor 6 which exchanges the data by the side of a drive. A sub** system consists of a central processing unit 8, a channel unit 9, an external memory control unit 10, and external storage 11 that connects with this similarly. Moreover, the external memory control unit 3 and the external memory control unit 10 are connected by the interface cable 12. When a system is comparatively built by this interface cable in a remote place, a fiber optic cable etc. is mentioned to it.

[0012] Drawing 2 shows the information processing system shown in drawing 1 further to details. It is asynchronous and the channel-command analysis section 14 and the command issuance section 15 are formed in the updating opportunity managed table 13 and the channel processing activation processor 5 which manage the information on whether doubleness of the data of a positive side system and a sub** system is performed to the external memory control unit 3 [other than numbers 4-6], or it carries out. Moreover, the service processor (SVP) 16 which can set up the renewal opportunity of data to the external storage 11 of a sub** system (a synchronization or asynchronous) etc. is connected to the external memory control unit 3 of a positive side system.

[0013] Drawing 3 and drawing 4 are examples of the updating opportunity managed table 13 which manage the renewal opportunity of data. Drawing 3 has been the truck processing information B which shows having set up a synchronization and asynchronous by the device processing information A which shows having set up a synchronization and asynchronous per device, and the information storage unit (here truck unit) of arbitration. Drawing 4 serves as the bit map table D showing a synchronization in the bit map table C showing a synchronization in a device unit, and an asynchronous established state, and a truck unit, and an asynchronous established state. The thing whose bit is ON and for which it is asynchronous and renewal of data to a sub** system is performed however is shown.

[0014] When a user sets up a synchronization and asynchronous in a device unit from SVP16, this information is written in the device processing information A, and then the bit of the device number of the bit map table C concerned is turned on. The time of being in the volume pair condition that the time of a system construction and external storage 7 and 11 have the same data, as an opportunity which sets up this information etc. can be considered. Moreover, when a user sets up a synchronization and asynchronous in a truck unit from SVP16, the information is first written in the truck processing information B on drawing 3 . Next, the bit of the bit map table D of the range where the specified truck corresponds is turned on.

[0015] It is thought that doubleness of data has a high possibility of being set up so that it may carry out synchronously since the consistency of the data of a positive side system and a sub** system is thought as important when setting out of such information is renewal of data with a significance high for a user, and a possibility of being set up so that it may carry out by being conversely asynchronous in renewal of data with which a response is thought as important although significance is not so high is high.

[0016] Drawing 5 is a flow chart which shows an example of an operation of processing of this operation. Hereafter, it explains according to this flow chart.

[0017] If the external memory control unit 3 of a positive side system receives a command from a central processing unit 1 (step 101), the light command with which the command performs renewal of data will distinguish whether it is no (step 102). When it is judged that it is not a light command here, the usual processing to the command is performed (step 103), and it returns to step 101. When it distinguishes from a light command at step 102, with reference to the device processing information A, it distinguishes whether it is set up per device (step 104).

[0018] When set up per device, with reference to the bit map table C, the bit of a device number with a demand sees whether it is ON (step 105). When the corresponding bit is ON, the renewal of data to a sub** system judges that it carries out by being asynchronous, and only the external storage of a positive side system updates data (step 108), and it returns to step 101. At this time, a channel end and the device end are reported to a central processing unit, and updating data is transmitted to the external memory control unit of a sub** system after that.

[0019] When the bit of a demand device number is OFF at step 105, it judges that renewal of data of a sub** system is performed synchronizing with the time of command receipt, and with the renewal of data of the external storage of a positive side system, updating data is transmitted to the external memory control unit of a sub** system (step 109), and it returns to step 101. At this time, a channel end is reported to a central processing unit, and the device end is reported, after data transfer completing to the external memory control unit of a sub** system.

[0020] When it is not setting out in a device unit at step 104, with reference to the truck processing information B, it distinguishes whether it is setting out in a truck unit (step 106). When it is setting out in a truck unit, with reference to the bit map table D, the bit of a truck range with a demand sees whether it is ON (step 107). When the corresponding bit is ON, it judges that renewal of data to a sub** system is performed in asynchronous, and only the external storage of a positive side system updates data (step 108), and returns to step 101.

[0021] When the bit of the truck range which had the demand at step 107 is OFF, it judges that renewal of data of a sub** system is performed synchronizing with the time of command receipt, and with the renewal of data of the external storage of a positive side system, updating data is transmitted to the external memory control unit of a sub** system (step 109), and it returns to step 101.

[0022] In not being setting out in a truck unit at step 106, either, it judges that renewal of data of a sub** system is performed synchronizing with the time of command receipt, and with the renewal of data of the external storage of a positive side system, updating data is transmitted to the external memory control unit of a sub** system (step 109), and it returns to step 101.

[0023]

[Effect of the Invention] According to this invention, it is asynchronous and it is possible to choose whether according to the property and significance of data to double, doubleness of the data of a positive side system and a sub** system is performed synchronously or it carries out, and to double data efficiently.

[0024] Moreover, it is possible to realize multiplexing of reliable data, suppressing lowering of the original information processing capacity which a system has to the minimum.

[Translation done.]

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平11-338647

(43) 公開日 平成11年(1999)12月10日

(51) Int.Cl.⁶

G 0 6 F 3/06

識別記号

3 0 4

P I

G 0 6 F 3/06

3 0 4 E

審査請求 未請求 請求項の数 2 O L (全 6 頁)

(21) 出願番号 特願平10-142571

(22) 出願日 平成10年(1998)5月25日

(71) 出願人 000005108

株式会社日立製作所
東京都千代田区神田駿河台四丁目6番地

(72) 発明者 水主 和人

神奈川県小田原市国府津2880番地 株式会
社日立製作所ストレージシステム事業部内

(72) 発明者 長瀬 光男

神奈川県小田原市国府津2880番地 株式会
社日立製作所ストレージシステム事業部内

(72) 発明者 木村 森久

神奈川県小田原市国府津2880番地 株式会
社日立製作所ストレージシステム事業部内

(74) 代理人 弁理士 小川 勝男

最終頁に続く

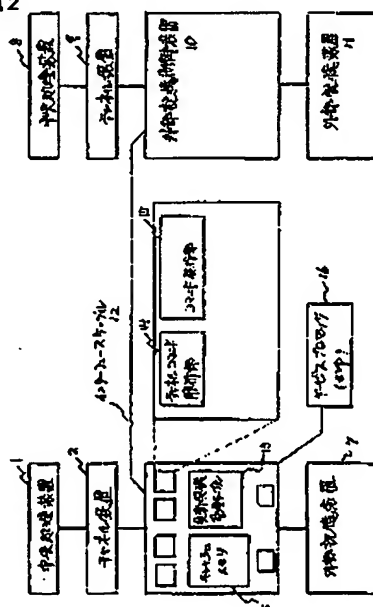
(54) 【発明の名称】 二重化データの形成方法

(57) 【要約】

【課題】本発明の目的は、二重化するデータの特性や重要度に応じて正側システムと副側システムとのデータの二重化を同期で行うか、非同期で行うかを選択して効率的にデータの二重化を行う技術を提供することにある。

【解決手段】正側システムの外部記憶制御装置3が上位装置よりライトコマンドを受領すると、要求のあったデバイス、或いはトラックのバットマップC、Dを参照しデータの二重化の契機を判別したのち、同期、或いは非同期にて副側システムのデータ更新を行う。

図2



(2)

特開平11-338647

1

【特許請求の範囲】

【請求項1】 外部記憶装置とこれを制御する外部記憶制御装置（以下、これらを総称して正側システムとする）と、上位装置から構成されるデータ処理システムであり、任意の情報通信リンクで接続された別の外部記憶制御装置配下の外部記憶装置（以下、これらを総称して副側システムとする）において同一のデータをバックアップデータとして保持するデータ二重化システムにおいて、副側システムへの情報の更新をデータの単位ごとに同期、或いは非同期に行えることを特徴とする二重化データの形成方式。

【請求項2】 請求項1において、副側システムの情報の更新はデータの単位ごとに同期、或いは非同期で行うことを任意に設定可能とすることを特徴とする二重化データの形成方式。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は、二重化データの形成方式及び外部記憶制御装置に関し、詳細には、正側システムの外部記憶装置内と副側システムの外部記憶装置内とに同一のデータを保持しデータの二重化を図ること、データの信頼性、安全性を向上させる技術に関し、特に、データ転送技術に関するものである。

【0002】

【従来の技術】 上位装置で使用するデータを外部記憶装置に保持し、データの二重化を行う方法として同一データを持つ外部記憶装置を異なる外部記憶制御装置配下に設け、その各々の外部記憶制御装置間を任意の情報通信リンクで接続することが一般的である。

【0003】 公知例特表平8-509565号公報にある様に、このようなシステム環境下で上位装置から書き込み命令が発行された場合に、一次データ記憶システムの外部記憶制御装置がこれと接続されている二次データ記憶システムの外部記憶制御装置に対し同様な書き込み命令を発行することでデータ二重化を図る技術がある。具体的には、上位装置からの書き込み命令受領時と同期して二次データ記憶システムの外部記憶制御装置に同様の書き込み命令を発行する技術で、この方法では二次データ記憶システムのデータ更新が終了するまで次の書き込み命令は受領されないで、確実にデータの二重化が実現できる。しかし、二次データ記憶システムの書き込み処理が完了するまで上位装置への完了報告が待たされる為、上位装置へのレスポンスは大きく効率が悪いシステムとは言えない。

【0004】 又、このようなシステム環境は、地震、火災等の災害によるデータ喪失を最小限に抑えることを目的に比較的遠隔地に設置される可能性が高い為、システムのスループットにも大きな影響を及ぼしてしまう。

【0005】 これに対し、上位装置からの書き込み命令受領時とは非同期に二次データ記憶システムの外部記憶

2

制御装置に書き込み命令を発行する技術がある。この方法では二次データ記憶システムの外部記憶制御装置への書き込み命令発行が上位装置とのデータの享受に関わらず行える為、レスポンスタイムは早くなるが、非同期に二次データ記憶システムのデータ更新が行われる為、一次データ記憶システムと二次データ記憶システムとのデータの整合性がとれない時間帯が存在する。この為、災害時のデータのリカバリが複雑になる。

【0006】 更に、上記の二つの手段は、あらかじめどちらか一方を選択しなければならない。つまり、非同期処理で更新を行っても良いようなデータをあえて同期処理にて行わなければならない、といったような非効率的な処理が発生してしまう。

【0007】

【発明が解決しようとする課題】 本発明の目的は、二重化するデータの特性や重要度に応じて、正側システムと副側システムとのデータの二重化を同期して行うか非同期で行うかを選択して効率的にデータの二重化を行う技術を提供することにある。

【0008】 本発明の他の目的は、システムが持つ本来の情報処理能力の低下を最小限に抑えつつ、信頼性の高いデータの二重化を実現する技術を提供することにある。

【0009】

【課題を解決するための手段】 本発明の目的は、正側システムの外部記憶制御装置が

（1）同期、非同期を認識する手段と、

（2）上位装置、或いはサービスプロセッサを介して、任意の記憶デバイス、情報記憶単位で正側システムと副側システムとのデータの二重化を同期して行うか、非同期で行うかを設定する手段とで達成される。

【0010】

【発明の実施の形態】 以下、本発明の一実施例の形態を図面を参照しながら説明する。

【0011】 図1は本発明の一実施例の形態であるデータ処理システムの構成図である。正側システムには中央処理装置1、チャネル装置2、外部記憶制御装置3とこれに接続される外部記憶装置7から構成される。又、外部記憶制御装置3は、チャネル装置2とのデータのやりとりを行うチャネル処理実行プロセッサ4、キャッシュメモリ5、及びドライブ側とのデータのやりとりを行うドライブ処理実行プロセッサ6を内蔵している。同様に副側システムも中央処理装置8、チャネル装置9、外部記憶制御装置10、これに接続される外部記憶装置11で構成される。又、外部記憶制御装置3と外部記憶制御装置10はインターフェースケーブル12で接続されている。このインターフェースケーブルには、比較的遠隔地にシステムが構築される場合、光ファイバケーブル等が挙げられる。

(3)

待開平11-338647

3

【0012】図2は、図1に示された情報処理システムをさらに詳細に示したものである。外部記憶制御装置3には番号4～6の他に、正側システムと副側システムとのデータの二重化を同期して行うか、非同期で行うかの情報を管理する更新契機管理テーブル13、チャンネル処理実行プロセッサ5にはチャンネルコマンド解析部14、コマンド発行部15が設けられている。又、正側システムの外部記憶制御装置3には、副側システムの外部記憶装置11へのデータ更新契機（同期、又は非同期）等の設定を行えるサービスプロセッサ（SVP）16が接続

されている。
【0013】図3、及び図4は、データ更新契機を管理する更新契機管理テーブル13の一側である。図3は、デバイス単位で同期、非同期を設定したことを示すデバイス処理情報A、任意の情報記憶単位（ここではトラック単位）で同期、非同期を設定したことを示すトラック処理情報Bとなっている。図4は、デバイス単位での同期、非同期設定状態を示すビットマップテーブルC、トラック単位での同期、非同期設定状態を示すビットマップテーブルDとなっている。ビットがONのところが副側システムへのデータ更新を非同期で行うことを示している。

【0014】ユーザがSVP16からデバイス単位で同期、非同期を設定した場合、この情報はデバイス処理情報Aに書き込まれ、次にビットマップテーブルCの当該デバイス番号のビットがONされる。この情報を設定する契機としては、システム構築時や外部記憶装置7と11とが同一データを持つボリュームペア状態である時、等が考えられる。又、ユーザがSVP16からトラック単位で同期、非同期を設定した場合には、まず図3のトラック処理情報Bにその情報が書き込まれる。次に、指定されたトラックの該当する範囲のビットマップテーブルDのビットがONされる。

【0015】これらの情報の設定は、ユーザにとって重要度の高いデータの更新であるときには正側システムと副側システムとのデータの整合性が重視される為、データの二重化は同期して行うように設定される可能性が高く、逆に重要度はそれ程高くないがレスポンスが重視されるようなデータの更新では、非同期で行うように設定される可能性が高いと考えられる。

【0016】図5は、本実施の処理形態の作用の一例を示すフローチャートである。以下、このフローチャートに従い説明する。

【0017】正側システムの外部記憶制御装置3は中央処理装置1からコマンドを受領すると（ステップ101）、そのコマンドがデータ更新を行うライトコマンドが否かを判別する（ステップ102）。ここでライトコマンドでないと判断した場合は、そのコマンドに対する通常の処理を行い（ステップ103）、ステップ101に戻る。ステップ102でライトコマンドと判別した場

4

合、デバイス処理情報Aを参照し、デバイス単位で設定されているかどうかを判別する（ステップ104）。

【0018】デバイス単位で設定されていた場合は、ビットマップテーブルCを参照し、要求のあったデバイス番号のビットがONかどうかをみる（ステップ105）。該当するビットがONの場合は、副側システムへのデータ更新は非同期で行うことと判断し、正側システムの外部記憶装置のみデータの更新を行い（ステップ108）、ステップ101に戻る。この時、中央処理装置にはチャンネルエンド、デバイスエンドを報告し、その後、副側システムの外部記憶制御装置へ更新データを転送する。

【0019】ステップ105で要求デバイス番号のビットがOFFであった場合には、コマンド受領時と同期して副側システムのデータ更新を行うことと判断し、正側システムの外部記憶装置のデータ更新と共に、副側システムの外部記憶制御装置へ更新データを転送し（ステップ109）、ステップ101に戻る。この時、中央処理装置にはチャンネルエンドを報告し、副側システムの外部記憶制御装置へのデータ転送完了後、デバイスエンドを報告する。

【0020】ステップ104でデバイス単位での設定ではない場合、トラック処理情報Bを参照し、トラック単位での設定であるかどうかを判別する（ステップ106）。トラック単位での設定であった場合は、ビットマップテーブルDを参照し、要求のあったトラック範囲のビットがONかどうかをみる（ステップ107）。該当するビットがONの場合は、副側システムへのデータ更新は非同期にて行うことと判断し、正側システムの外部記憶装置のみデータの更新を行い（ステップ108）、ステップ101に戻る。

【0021】ステップ107で要求のあったトラック範囲のビットがOFFであった場合は、コマンド受領時と同期して副側システムのデータ更新を行うことと判断し、正側システムの外部記憶装置のデータ更新と共に、副側システムの外部記憶制御装置へ更新データを転送し（ステップ109）、ステップ101に戻る。

【0022】ステップ106でトラック単位での設定でもない場合には、コマンド受領時と同期して副側システムのデータ更新を行うことと判断し、正側システムの外部記憶装置のデータ更新と共に、副側システムの外部記憶制御装置へ更新データを転送し（ステップ109）、ステップ101に戻る。

【0023】

【発明の効果】本発明によれば、二重化するデータの特性や重要度に応じて、正側システムと副側システムとのデータの二重化を同期して行うか、非同期で行うかを選択して効率的にデータの二重化を行うことが可能である。

【0024】又、システムが持つ本来の情報処理能力の

(4)

特開平11-338647

5

低下を最小限に抑えつつ信頼性の高いデータの多重化を
実現することが可能である。

【図面の簡単な説明】

【図1】本発明の一実施例におけるデータ二重化システ
ムを構成する情報処理システムの構成図である。

【図2】図1に例示された情報処理システムの構成をさ
らに詳細に示した内部構成図である。

【図3】本発明の一実施例の形態である各種制御情報の
一例を示す概念図である。

【図4】本発明の一実施例の形態である各種制御情報の
一例を示す概念図である。

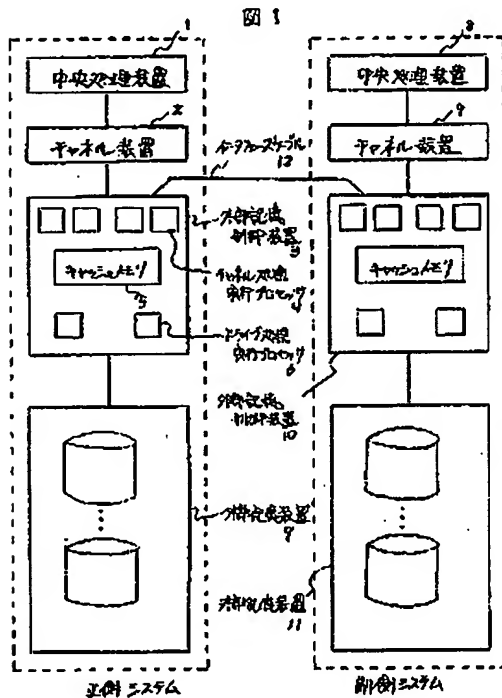
【図5】本発明の一実施例の形態である情報処理システ

* ムの作用のフローチャートである。

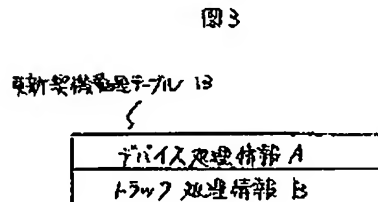
【符号の説明】

1…中央処理装置、2…チャネル装置、3…外部記憶制
御装置、4…チャネル処理実行プロセッサ、5…チャ
ネル処理実行プロセッサ、6…ドライブ処理実行プロセ
ッサ、7…外部記憶装置、8…中央処理装置、9…チャ
ネル装置、10…外部記憶制御装置、11…
外部記憶装置、12…インターフェースケーブル、13
…更新契機管理テーブル、14…チャネルコマンド解析
部、15…コマンド発行部、16…サービスプロセッ
サ。

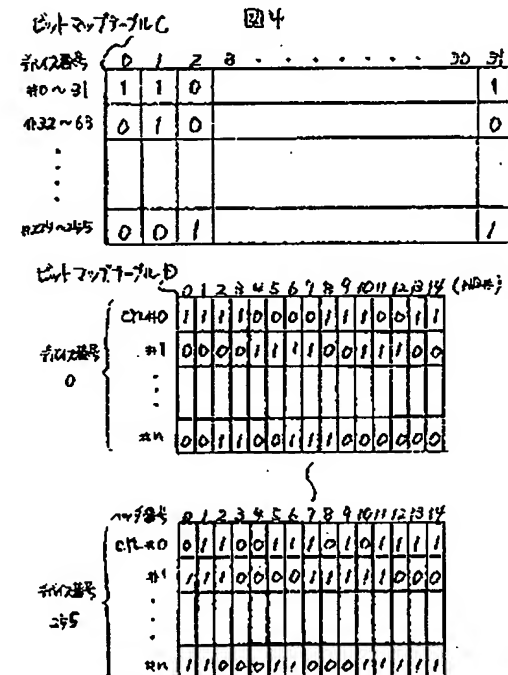
【図1】



【図3】



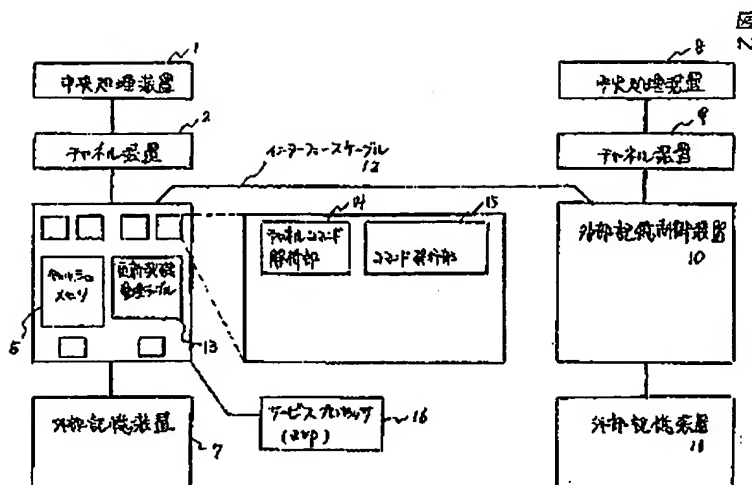
【図4】



(5)

特開平11-338647

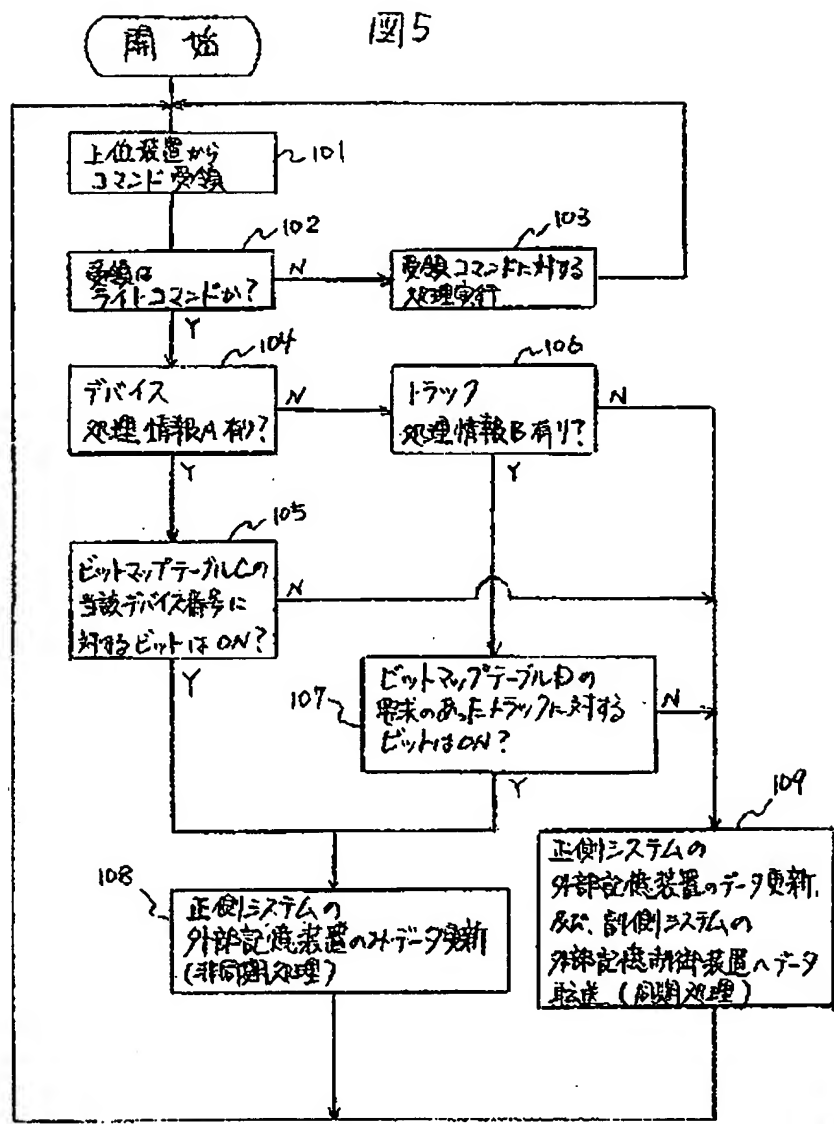
【図2】



(5)

特開平11-338647

【図5】



フロントページの続き

(72)発明者 小出 雄

 神奈川県小田原市国府津2880番地 株式会社
 日立製作所ストレージシステム事業部内